

Carbohydrate utilization capacity of uncyanobacterial consortia derived from hypersaline lake microbial mats

Semen A. Leyn¹, Yukari Maezato², Margie Romine² and Dmitry A. Rodionov^{1,3}

¹*Institute for Information Transmission Problems RAS (Kharkevich Institute), Moscow, Russia;*

²*Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA;*

³*Sanford-Burnham Medical Research Institute, La Jolla, CA, USA.*

semen.ley@gmail.com ; rodionov@burnham.org

Hot Lake is a meromictic, hypersaline lake located in north of Washington state, USA. This lake harbors a phototrophic microbial mat community that is characterized by its stability in dramatic seasonal environmental changes. The complexity of natural mat community hinders to analyze metabolic interactions between its members. To manage this obstacle, several uncyanobacterial consortia (UCC) were sampled from natural mat community and sequenced. UCC is a microbial community in which one cyanobacterium serves as the sole primary producer that supplies autotrophically-derived carbon to several heterotrophic bacteria. In mutualistic interactions, the heterotrophs can provide to cyanobacteria key metabolites and help to scavenge wastes. Next-generation sequencing and metabolic reconstruction techniques permit a broad characterization of consortial membership and provide the basis for studying metabolic interactions between the community members.

Recently available metagenomes for two UCC (Oscar and Ana) revealed a shared set of 19 microbial operational taxonomic units (OTU) that correspond to 17 distinct species of heterotrophic bacteria and two cyanobacterial species, whose individual genomes were reconstructed by co-assembly and abundance profile binning. The identified UCC heterotrophs mostly belong to Alphaproteobacteria (10 OTU), Gammaproteobacteria (5 OTU) and Bacteroidetes (2 OTU), and their reconstructed genomes range between 2.7 and 4.3 Mb [1]. We used this genomic data and a comparative genomics approach to infer carbohydrate utilization pathways and their transcription regulation in these UCC bacteria.

First we scanned all UCC protein sequences against the KEGG Orthology, Pfam, GO and SEED databases to reveal genes potentially involved in sugar utilization. Then we analyzed the genomic and functional contexts of gene loci encoding the obtained proteins and reconstructed the respective metabolic pathways and transcriptional regulons. For the

comparative genomics-enabled pathway and regulon inference, we used additional closely-related bacterial genomes available in the Integrated Microbial Genomes (IMG) database of the Joint Genomic Institute.

As result, we observed catabolic pathways for more than 20 various carbohydrates and their derivatives that can serve potential carbon sources for the analyzed heterotrophic bacteria in UCC. These carbon sources include glucose, galactose, N-acetylglucosamine, N-acetylgalactosamine, rhamnose, fucose, xylose, fructose, sorbitol, mannitol and other. Distribution of the inferred catabolic pathways has highly mosaic structure across the UCC genomes. The largest number of pathways was observed in two representatives of the *Halomonas* genus (13 pathways), one *Rhodobacteriaceae* sp. (12 pathways) and one *Algoriphagus* sp. (11 pathways). In other UCC genomes, we found up to seven carbohydrate utilization pathways per genome. The most widely distributed pathway in UCC was the alpha-glucosides utilization pathway that is present in 11 genomes. The DeLey-Doudoroff pathway of galactose utilization was found in 5 Alphaproteobacterial members. The catabolic pathways for sugar alcohols (sorbitol, mannitol) were found in 4 Rhodobacteriaceae and 2 *Halomonas* spp. genomes, however they use different sets of genes for these pathways. Other sugar catabolic pathways are less abundant among the UCC members as they were found in 1-3 genomes. The reconstructed catabolic pathways were partially validated by phenotypic analyses of UCC community members that were isolated in a pure culture, and their abilities to utilize an individual sugar as a sole source of carbon and energy was further tested.

In the Rhodobacteriaceae sp., we found a novel catabolic pathway that was proposed to be involved in mannoheptulose utilization. The pathway consists of an ABC transporter, two kinases and an aldolase from the tagatose 1,6-diphosphate aldolase family. The sugar-binding component of the ABC transporter has 83% similarity to an ABC transporter from *Agrobacterium vitis* that was known to bind mannoheptulose (Steve Almo & Matt Vetting, personal communication). Mannoheptulose is a heptose that has a similar structure to tagatose. To our knowledge, the proposed mannoheptulose catabolic pathway is the first case of a pathway for heptose utilization and will require further experimental validation.

The identified pathways are potentially regulated by local transcription factors (TFs) from the

IclR, LysR, SorC, OmpR, RpiR, ROK, DeoR, LacI, and GntR protein families, that are often co-localized with the respective sugar catabolic gene loci. The most populated families of sugar metabolism-associated regulators in UCC genomes are LacI and GntR. For the majority of the identified TFs we identified their cognate DNA binding motifs and reconstructed regulons using the bioinformatics approach. For *Halomonas* and *Rhodobacteriaceae* spp., we also reconstructed the central carbon metabolism transcriptional regulons, HexR and CceR respectively.

This research was supported by the Russian Science Foundation (grant #14-14-00289). Additional funding was provided by the Genomic Science Program (GSP), Office of Biological and Environmental Research (OBER), U.S. Department of Energy (DOE), and is a contribution of the Pacific Northwest National Laboratory (PNNL) Foundational Scientific Focus Area.

1. 1. J.K. Cole et al. (2014) Phototrophic biofilm assembly in microbial-mat-derived unicyanobacterial consortia: model systems for the study of autotroph-heterotroph interactions, *Front Microbiol.*, **5**:109.